SPECIFICATION

Please replace paragraphs [0001], [0079], [0082], and [0090] of the specification as follows:

[0001] "SYSTEM AND METHOD FOR PERFORMING ELECTRONIC INFORMATION RETRIEVAL USING KEYWORDS," Alain Franciosa et al., application Ser. No. ~~XX/XXX,XXX~~ 10/605,630, filed concurrently herewith and incorporated herein by reference. ~~(Docket No. D/A3358).~~

[0079] At 508, the inclusion ratio "I", which represents the percentage (e.g., a first percentage) of keyword weights in the input document that are in the document search results plus the tolerance ration T (as opposed to ratio "R" which is based on the number of keywords in the input document), is computed. The inclusion ratio "I" specifies one factor in the distance measurement. That is, an input document is considered an "inclusion" of search results if the input document is determined to be included in the search results (e.g., the input document is one page of a ten page search result document).

[0082] More specifically, the inclusion verification is done by computing an ordered inclusion ratio "I2" that is a function of an ordered sum of the weights of the keywords. That is, the inclusion ratio "I2" defines what percentage (e.g., a second percentage, a fourth percentage, etc.) of keyword weights along with a set of their neighboring keyword weights in the input document also exist in the search results document. The initial determination of an inclusion is not conclusive because the inclusion ratio "I" does not take into account the weights of the keywords coming from D2 as the inclusion ratio "I" is defined as the percentage defined by ratio of Sum3/Sum2, where:

[0090] At 512, the Jaccard similarity distance measure (e.g., a third percentage) is computed for the list of keywords D1 and D2. Through regression analysis of search results it has been determined that if the Jaccard similarity measure is greater than approximately ninety percent then it is likely that the document being evaluated is a

revision of the input document; otherwise, it is believed that the document being evaluated may be a document similar to the input document at which point the Jaccard similarity is returned as the measure of similarity S between keyword lists D1 and D2.